IMPLEMENTING
POWER BI PREMIUM CAPACITY

phData

CONTENTS

Choosing Power BI Premium for your organization is a big decision – one that can really enhance your organization's analytics capabilities. Moving forward with Power BI Premium is also an investment in your analytics infrastructure, so you want to make sure that it is implemented and scaled appropriately.

This whitepaper is dedicated to helping you understand the process of implementing Power BI Premium capacity as well as understanding some of its nuances for a successful deployment. Let's dive in!

WHAT IS POWER BI PREMIUM CAPACITY

Power BI Premium capacity is essentially a dedicated server space that you purchase from Microsoft that allows you to be the sole organizational user of a set amount of server capacity. You are not sharing server capacity with other organizations like you would be if you chose to use Power BI Pro licensing.

There are several different tiers of Premium capacity. The capacity model you select will designate the processing power behind your entire analytics workspace, and selecting the right one is crucial to the success and adoption of your analytics program regardless of the stage of your organization's analytics maturity.

Let's explore which capacity model might be right for you.

A SKU

SKUs (A1-A6) is a Platform-as-a-Service and set of APIs for Independent Software Vendors (ISVS) who develop and deliver applications to market. These ISVs choose to use Power BI as the data visualization layer of that application to add value to their own application.

Unfortunately, PowerBI assets contained in Power BI Embedded capacities cannot be accessed by a licensed Power BI user. They are only accessible by the customers of the IVS' applications:

- A1 ~$750/month
- A2 ~$1,500/month
- A3 ~$3,000/month
- A4 ~$6,000/month
- A5 ~$12,000/month
- A6 ~$24,000/month

P SKU

SKUs (P1 - P5) are enterprise features for embedding. The commitment is either yearly or monthly and includes a license to install PowerBI Report Server on-premises.

- P1 ~$5,000/month
- P2 ~$10,000/month
- P3 ~$20,000/month
- P4 ~$40,000/month
- P5 ~$80,000/month

EM SKU

SKUs (EM1-EM3) for organizational embedding. EM1 and EM2 SKUs are only available via volume licensing plans (Office 365) and you can't purchase them directly. In an environment with an EM SKU, we recommend choosing the PPU option for developers who are working in a separate workspace. This will allow them to avoid crossing hairs with business-critical reports that your free-licensed viewers will be interacting with.

- EM1 ~$625/month
- EM2 ~$1,250/month
- EM3 ~$2,500/month

The SKU represents the memory, storage, and set amount of resources consisting of processors where ALL SKUs are considered as capacity nodes. Each SKU contains operational limits on the number of DirectQuery and Live Connections per second, as well as the number of parallel model refreshes.

Here we will assume that you're planning on deploying Power BI for internal purposes and only focus on

the P SKUs. With P and EM SKUs, all viewers can leverage unlimited free licenses. You can also set your developers up with Pro or Premium Per User licenses.

The resources and limitations of each Premium SKU (and equivalently sized A SKU) are described below:

Capacity SKUs	Total v-cores	Backend v-cores	Frontend v-cores	Max memory per dataset (GB)[1,2,3]	DirectQuery/ Live connection (per second)[1,2]	Max memory per query (GB)[1,2,3]	Model refresh parallelism[2]
EM1/A1	1	0.5	0.5	3	3.75	1	5
EM2/A2	2	1	1	5	7.5	2	10
EM3/A3	4	2	2	10	15	2	20
P1/A4	8	4	4	25	30	6	40
P2/A5	16	8	8	50	60	6	80
P3/A6	32	16	16	100	120	10	160
P4/AY[4]	64	32	32	200	240	10	320
P5/A8[4]	128	64	64	400	480	10	640

[1] The Power BI Premium Utilization and Metrics app doesn't currently expose these metrics.

[2] These limits only apply to the datasets workload per capacity.

[3] The *Max memory per dataset (GB)* column represents an upper bound for the dataset size. However, some memory must be reserved for operations such as dataset refreshes and queries. The maximum dataset size permitted on a capacity may be smaller than the numbers in this column.

[4] SKUs greater than 100 GB aren't available in all regions. To request using these SKUs in regions where they're not available, contact your Microsoft account manager.

WHY CHOOSE POWER BI PREMIUM CAPACITY

There are many reasons why you might choose Power BI Premium Capacity, but there are a few specific scenarios that regularly drive organizations to Premium.

Organic Growth

In this scenario, adoption of Power BI often starts with a champion who uses the tool to develop solutions often in their team or a smaller ecosystem. Others in their organization take note and begin to use Power BI as well. Soon other teams take notice of the awesome analytics being produced by the champion's group and suddenly you have several teams using Power BI. In this scenario, most users will be on Power BI Pro licensing. With this licensing model, both contributors and viewers need to have a Pro license which costs $9.99 a month per user. With a P1 SKU of Premium being priced at $4,995 per month, having a user base of around 500 users makes the costs of Pro and a P1 Premium SKU about even.

Premium Features

As organizations scale their adoption of Power BI, there is a bigger chance that some of the features in Power BI Premium capacity become more applicable. You can find a list of some of these features here. Some of the common features that push people to Premium are larger data size limits, paginated reports, and deployment pipelines. These features are often essential for scaling your Power BI deployment across larger and more complex organizations.

Even if your situation does not align with either of these scenarios, Power BI Premium capacity could still be the right choice for you.

DEPLOYING POWER BI PREMIUM CAPACITY

When starting this process it is a good idea to loop in your Microsoft account manager to make sure they are aware of your decision to go with Power BI Premium. They will need to do some things on their end to make sure that you are able to procure the product you are looking for, so looping them in sooner than later is advisable.

CHOOSING THE RIGHT POWER BI PREMIUM SKU

Here are some questions that you should answer before choosing what Premium SKU is right for you and your organization.

What will this reporting environment be used for?

This is a very high level question, but it's one of the most important ones to understand. You need to understand the purpose of your Premium capacity environment to better comprehend the scope and scale of who will be using this capacity.

This question will also help prompt discussion around how many business groups and geographies will be using this instance - all things that you will need to know as you set up your capacity.

Approximately how many users will be using this capacity?

The number of users will generally be a function of how many business groups are using the capacity, but it's good to get a headcount so you can prepare for disseminating licenses to the appropriate people. At this time, it is also good to identify the folks who will be champions/developers for the reports so you can keep them engaged in the process.

What are the largest datasets that will be needed for reporting?

The answer to this question will help you determine how much RAM you need in your environment for refreshing datasets. When a Power BI imported dataset refreshes, Power BI keeps the original data model available for users while creating a temporary mirror model for the refresh.

This means that at the time of refresh, your dataset will essentially be twice the size, so make sure to keep that in consideration. You will want two numbers here:

1. What is the size of the largest data set for each business group?
2. What is the total size of data needing to be imported (preferably by dataset)?

These numbers will help guide you in your SKU selection.

Note: Make sure to ask about any future growth in the data sizes or scope so you can anticipate needing more capacity down the road. For example, currently the largest data set for Sales reporting is 1 GB, but that only includes one year of data and the East region. If they plan on adding two more years of data and three other regions, you need to extrapolate how large that future data set will be to get the most accurate prediction of capacity.

Another Note: Make sure to clarify how the current data is being modeled to understand if there might be differences between the current model and how it will be modeled in Power BI. If datasets are currently being modeled as a flat file then you might get some relief on data set size if you utilize a STAR schema.

How often will data need to be refreshed?

This will help you determine how many refreshes you need available for each imported dataset as Power BI Premium capacity has a limit of 48 scheduled refreshes a day while Power BI Pro licenses allow for eight refreshes a day.

This could also help you determine if you need to use Direct Queries instead of Import. Once you have the answers to these questions you will have a more clear understanding of what P SKU to use for your use case.

Let's walk through an example use case to showcase how one might decide on a P SKU given a set of fictional constraints.

- **What will this reporting environment be used for?** This reporting team has two distinct business groups needing to use Power BI Premium Capacity - Finance and Sales. There are no crossover users and their data should not be accessible to one another.

- **Approximately how many users will be using this capacity?** There will be approximately 500 viewers and 10 contributors for each group

- **What are the largest datasets that will be needed for reporting?** For Sales, the largest dataset is 15GB, but it only includes one year of data and they plan on expanding to two years of data. Their datasets add up to a total of 50GB. For Finance, the largest dataset will be 10GB and it includes all years required. All of their datasets add up to 15 GB.

- **How often will data need to be refreshed?** Sales will need their datasets to be refreshed 10 times a day while Finance will need refreshes on a daily basis.

Based upon the criteria given above, we would recommend a P2 SKU for this deployment. We know that they need Premium because of Sales' desire for 10 refreshes a day, and we know that the largest data set will be about 30 GB of data.
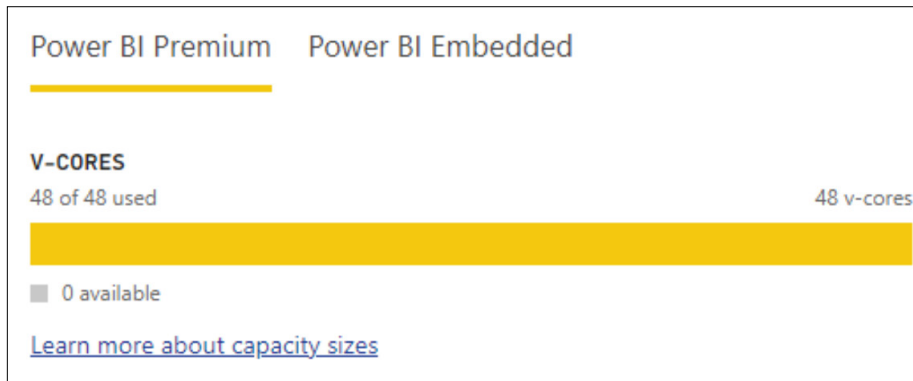
Although we generally suggest the largest data set x 2 for meeting the max size criteria, in this case we would suggest incremental refresh of their dataset. Once the first dataset is loaded, then we wouldn't be refreshing the entire dataset every day, just the new data. Also, because we can set up different workspaces for the two groups within the one capacity, we won't need separate capacities for the groups, hence the P2 recommendation.

Note: In this scenario, it would be very realistic to suggest a P3 SKU in case to anticipate further business groups being added to the dedicated capacity.

SETTING UP YOUR PREMIUM SKU

Once you have selected the desired Power BI SKU, make sure to align with your Microsoft representative so that they can start the process of getting the paperwork done to procure your dedicated capacity.

After the dedicated capacity has been assigned to your tenant, you can go to the Admin portal in Power BI Service and set up your new capacity. You should be able to see the nodes available at the top of the Capacity Settings section of your Admin portal.

Power BI Premium   Power BI Embedded

V-CORES
48 of 48 used   48 v-cores

☐ 0 available

Learn more about capacity sizes

If you have not yet assigned your capacities, go ahead and select the ***Set Up New Capacity*** button. This will enable you to assign your capacity pursuant to your organizational strategy. Keep in mind that the capacity that you purchase can be broken down into smaller capacities (you can assign a P3 SKU as three smaller P1 capacities).

Please note that the criteria we discussed when choosing the correct SKU, by choosing smaller SKUs, you will lower max data set size, etc.

Are you already using Azure AD groups for access?

If your organization is already using the O365 suite, then odds are you're using Azure Active Directory groups for access to a whole host of items, but you want to make sure before you start. If they are not, then you will need to set up your Azure Active Directory to be able to provision access once your workspaces are set up.

PARTITIONING WORKSPACES IN THE CAPACITY

Once you've set up the capacities, you will want to assign workspaces to the dedicated capacity. Navigate to the Admin Portal and then Capacity Settings. Click on the capacity that you want to add workspaces to and then select **Workspaces assigned to this capacity**. You can then select the **+ Assign workspaces** button at the top right.

From here you can assign workspaces based on the users, specific workspaces, or all of the organization's workspaces. If you have not created any workspaces yet, you will need to determine how to partition the workspaces appropriately first.

HOW MANY USER GROUPS DO YOU HAVE?

It's important that you first identify who your users are and how many different user groups you have. In this context, user groups refers to groups of people that require different views of data, or in some cases, different data entirely. Listed below are three examples of how to think about user groups.

- **Business Functions:** Are you catering to multiple business functions, i.e. HR, supply chain, marketing, finance, sales, etc.?
- **Business Teams:** Are you building reports for a single business function with multiple teams? For example, different order fulfillment methods within the supply chain business function, i.e. in-store pickup, curbside pickup, home delivery.
- **Location:** Is your team implementing solutions for a business function or business team that needs to see the data differently depending on geographical locations?

This is critical because it will either determine the amount of workspaces you provision, or the number of audience views you create within a single workspace app (or both).

IS ROW-LEVEL SECURITY (RLS) REQUIRED OR IS IT AN OPTION?

RLS is a strategy for implementing horizontal limitations on tables in a data model. To clarify, RLS limits the rows in a table an end-user can see based on specific criteria. For example: building a sales dashboard for regional sales managers.

Using RLS, you can limit the data the managers see to only include geographic locations relevant to them. This is important to understand and determine how many workspaces or audience views you will need to create.

Reference this blog post on how to set up row-level security in Power BI if your use case requires it.

WHAT ARE THE POWER BI PREMIUM WORKSPACE STRATEGIES?

Deciding on a workspace strategy is a crucial step in setting up a Premium Capacity. Leveraging the appropriate strategy to enable self service will result in the best user experience for developers and the report audience.

No matter the strategy you choose to use, there are a couple of best practices that should be adhered to. Let's cover these best practices and then also discuss the two recommended workspace strategies.

Traditional Development Lifecycle

A traditional development lifecycle can be accomplished by creating separate workspaces for development, test, and production environments. The development workspace is used for ongoing development of existing datasets and reports or net new content. Developers and admins should be granted access to this workspace.

The test workspace is where UAT testers can perform data validations and dashboard functionality testing. Admins and developers will need to be granted access to this workspace. UAT testers should be granted access to the workspace app for functionality testing and/or access to the datasets for data validation, especially if end users will be given access to the underlying dataset for self service.

The production workspace is strictly for production-ready datasets and reports or workspace apps that your end users will be utilizing. Developers and admins should be granted access to the production workspace. Ideally end users will be granted access to the workspace app with the option to turn on access to the underlying dataset if that level of self service is needed.

Power BI Deployment Pipelines

Deployment Pipelines are designed to be a reusable and efficient pipeline automating the movement of Power BI content through the development, test, and production workspaces. Access to deployment pipelines is separate from access to workspaces and ideally should be limited to a select group of people, like workspace admins and lead developers.

Learn more about Power BI Deployment Pipelines and how to set them up.

CREATING WORKSPACES

To create a workspace, navigate to the left side of Power BI Service and select the Workspaces ribbon. At the bottom you'll find a yellow button labeled **Create a workspace**. Select this, name your workspace, and complete the other optional fields, if desired.

POWER BI LICENSING

Although we are setting up Power BI Premium capacity, it is still necessary to provision licenses for all users, to either contribute or view content depending on their role. This means that for users to view content they will still need to be provisioned Power BI Free licenses and contributors will need either Pro or Premium Per User (PPU) licenses to publish content.

Let's discuss these license types to help you determine which ones you'll need for your deployment.

LICENSE TYPES

There are three types of Power BI per-user licenses: Free, Pro, and Premium Per User. Determining which kind of license a user needs is dependent on where content is stored, how the user will interact with the content, and if that content needs to utilize any Premium features.

Content storage itself will depend on your organization's subscription license type. Power BI Premium is the most commonly-used type of organizational subscription, which is a capacity-based license. Using Premium allows free license users to interact with content in workspaces, whereas with a non-premium capacity subscription, these users can only create content in their own private workspace without the ability to share it.

Power BI organizations - with free and Pro per-user licenses - will have a shared and limited ability to process content. Pro users will be able to collaborate with other pro users, but unfortunately not with free users.

License Type Capabilities

Features	Power BI Free	Power BI Pro	Power BI Premium
Space allocation	1GB	10GB *	100 TB
Power BI Desktop	✓	✓	✓
Development features	✓	✓	✓
DirectQuerv	-	✓	✓
Gateways	-	✓	✓
Row- level security	-	✓	✓
Content pack	-	✓	✓
Work spaces	-	✓	✓
Data streaming	up to 10K rows per hours	unlimited	unlimited
Dataset refresh frequency	up to once a day	up to 8 times a day	Up to 48 times a day

POWER BI PRO VS. POWER BI PREMIUM PER USER

What Does Power BI Pro Include?

As shown above, Power BI Pro is the full version of Power BI. It comes with the ability to develop reports and dashboards, as well as unlimited sharing and consumption of your reports and those shared by other users.

The Pro license allows you a maximum dataset size of 1GB and maximum storage of 10GB per user. Data can be refreshed up to eight times a day and costs $10/month/user if your organization isn't on an Microsoft 365

E5 license, where it's already included. Power BI Pro is suitable for smaller organizations, or instances where only a small number of users need to be able to create and collaborate on reports.

WHAT DOES POWER BI PREMIUM PER USER INCLUDE?

Power BI Premium Per User is a fairly new license type from Microsoft that gives organizations the ability to take advantage of Premium capacity features on a per-user basis. It includes all Power BI Pro license capabilities, along with features like paginated reports, AI, and others previously only available with Premium capacity.

As a Premium Per User licensed user you receive an increased model size limit of 100GB, up to 48 refreshes each day, and maximum storage of 100TB (just like Premium capacity). Premium Per User costs $20/month/user and includes almost every feature available in Premium capacity, making it a smart option for smaller organizations that still need the robust capabilities offered in Premium capacity.

Note: for PPU licensing is that viewers of reports created in PPU workspaces will still need PPU licenses to view. The same thing applies to reports created using Pro licenses - Pro licenses are required for viewers. When you upgrade to Premium capacity, you are able to leverage free licenses for viewers of your content in Premium capacity workspaces.

PROVISIONING GROUP BASED LICENSING

At phData we recommend setting up Azure Active Directory groups for group-based license provisioning. This recommendation is based on the assumption that your Power BI deployment is large enough to justify it.

Similar to workspaces, having an AD group that grants you a Power BI license based upon your inclusion in the AD group streamlines obtaining a license into the AD process.

To set up your group based licensing, make sure that you have the appropriate AD group set up first. You'll want the AD group name to indicate the license type and any other breakdown details in the licensing that your organization may have (team, region, organization, etc.).

Next go to your Azure Portal and select the **Azure Active Directory** application and select **Licenses** from the left-hand side. Then select **All Products** and go down to your **Power BI licenses**. It is important to remember that the licenses here will just be your license-based products, not Power BI Premium.

Your Power BI Premium SKUs will be listed in your **Power BI Admin Portal**. Choose the license you want to link to your licensing AD group and then select **Licensed Groups** on the left-hand side. When adding a group to this license, select the **Assign** button in the top left and then add the appropriate group.

Now, when people are granted access to the licensing AD group, they will be issued the licenses based upon this group based licensing.

Home > phdata.io | Licenses > Licenses | All products > Power BI Premium Per User

Power BI Premium Per User | Licensed groups ...

+ Assign   × Remove license   ⟳ Refresh | ☰☰ Columns | ⧸ Got feedback?

General

👤 Licensed users

👥 Licensed groups

👤 Service plan details

Name

Search by name

Name

There are no groups with licenses assigned for this product.

PROVISIONING ACCESS TO WORKSPACES

Once you have sorted out your workspaces, you will need to ensure that you are provisioning access to your datasets adequately. Having a detailed plan for provisioning Power BI licenses and Power BI Service workspace access in your organization will assist in keeping your deployment functionality efficient and smooth.

PROVISIONING LICENSES & ACCESS IS EASY, SO WHY ARE WE TALKING ABOUT THIS?

Although Power BI makes allocating access to data, reports, and workspaces fairly simple, we recommend some procedures for future-proofing your Power BI deployment.

Provisioning access seems simple at first, but almost every organization has their own approach to accomplishing this task. Data governance and security implications warrant an intentional approach to your provisioning.

HOW TO PROVISION ACCESS TO WORKSPACES

If you are deploying Power BI to your organization for the first time, you'll be relieved to know that the Azure Active Directory (AD) groups that you have already set up for your other Office applications can also be used out of the box with Power BI.

We recommend using these same AD groups as the source of access rather than granting access to content in the Power BI service. Essentially, you are just controlling access to your workspaces through your AD group process. Doing this will make your Power BI access process much more efficient and simple.

Click on the **Access** icon in the top right of any workspace to grant access to a workspace. You'll see this pop-out screen. Here you can enter either the name of someone in your Active Directory or an Azure AD group name in the **Enter email address** field.

Then you select the workspace role that the person or group should have and click **Add**. If you aren't clear on the different roles, learn more about workspace roles.

That was pretty simple, right? Well, it can get a little more complicated from there. In the example just showcased, we had an Admin and then one group gaining access to the workspace as Members. This example is very straightforward and doesn't take into account a large deployment that contains multiple workspaces and varying roles.

LARGER WORKSPACE DEPLOYMENT SCENARIO

The access gets a little more complicated in situations where you have a lot of different workspaces with different groups taking different roles within the workspaces. In these scenarios, our suggestion is to create different Azure AD groups for each of the roles that you expect users to need for each of the workspaces.

If a deployment of Power BI will have five different workspaces with each workspace having needs for both

Contributor and Member roles for example, our suggestion would be to create a total of 10 different Active Directory groups in your Azure Portal like this:

1. Workspace1-Contributor
2. Workspace1-Member
3. Workspace2-Contributor
4. Workspace2-Member
5. Workspace3-Contributor
6. Workspace3-Member
7. Workspace4-Contributor
8. Workspace4-Member
9. Workspace5-Contributor
10. Workspace5-Member

Once your workspace is set up, go in and issue these groups access to the workspace so your access is then controlled through the AD group request process and not subject to ad hoc requests to an actual workspace in Power BI Service.

For those new to creating AD groups, simply go to your *Azure Portal* and select the *Azure Active Directory* application. Next, select *Groups* and then *New Group* in the top left.

EXCEPTIONS TO THIS ACCESS PROCESS

Maybe you're reading this post and you say to yourself, this doesn't fit my organization's requirements - well you might be right! Here are some common exceptions:

Testing Workspaces

During UAT, your testing group may be very ad hoc and smaller in number. In this instance, you can grant access to a report without granting access to the underlying data by granting testers Direct Access to the report itself.

You can do this by clicking on the vertical ellipses next to a report in a workspace and then clicking *Manage Permissions*. Select *Direct Access* and then *Add User* to allow testers access to the appropriate report.

Access is Controlled via Workspace App

This scenario is less of an exception than a deviation from thenorm. If you are wanting to grant access to just the app (which is a best practice for viewers that we recommend) then you will grant access to the viewer AD group in the app instead of the workspace.

This ensures that the end viewers are only seeing the final product in the app instead of the workspace itself. Learn more about workspace apps.

Dev Workspaces

Although we still highly recommend creating an AD group for developers that will have access to the Dev workspaces, you may only want to create one group for Dev access since it will be a select group that is actually developing the content for publishing.

If your organization or group of developers is small enough, your Power BI admin may want to grant access to Dev workspaces on an individual basis. This is understandable due to the flexibility and quickness that it provides the Admin to control who has access to the Dev workspace.

CHOOSING AZURE VIRTUAL MACHINES FOR GATEWAYS

When deploying Power BI, you'll want to decide whether or not you should use an on-premise gateway to connect to your data. As suggested by the name, an on-premise gateway will be needed to authenticate to an on-premise data source, but they can also be used to improve continuity and security with cloud to cloud connections.

WHAT IS AN ON-PREMISES GATEWAY?

An on-premises gateway is essentially a link between two data sources, which is generally from an on-premises data source to the cloud allowing for a secure transfer of data between the two locations. The gateway itself is a Microsoft program that you download and configure based on your desired connection. It is maintained by a select group of gateway admins. These gateways allow for secure connections for all of the data sources you connect through the gateways in your reporting environment.

Although you can also host gateways on your local desktop machine, it is recommended – and preferable – to host them on Virtual Machines (VMs) for a few reasons:

1. The VM you select will likely be more powerful than your desktop, which means that it will be able to better handle the loads of queries.

2. The safety of the gateway does not rely on physical hardware that could easily be lost or destroyed.

3. If the owner of the local desktop machine housing the gateway were to leave the company, you would have to set it up all over again on a different machine. With a VM, you can simply transfer credentials – and maintain governance – with a service account.

WHAT IS A VIRTUAL MACHINE?

A VM is essentially a dedicated space – or capacity – on a physical server that houses all of the components of a physical computer like CPU, disks, and memory. Think of it not as a physical desktop computer, but rather a virtualized version of a computer that exists purely as code. This concept is important to understand because VMs are where you will host your Power BI Gateways rather than hosting them on your actual physical machine.

While VMs can be scaled up or down easily to meet the needs of changing workloads, physical machines require you to replace the entire physical computer if workloads change dramatically. There are a multitude of types of VMs available for purchase on your Azure Portal and each have different capabilities and purposes. Choosing the right VM for your needs is imperative for the long term success of your Power BI deployment.

HOW TO CHOOSE THE RIGHT VIRTUAL MACHINE

Before diving into the different types of VMs, we should first understand the use case for the VMs. These VMs will be handling the workloads for your data refreshes and/or queries to your data sources.

An important consideration will be whether developers will be utilizing Direct Query or Import mode for their data model. Also, you will want to know how large the datasets are that you will be querying. Our recommendation for VM sizing and type will be based on this information.

Once you've gathered the aforementioned information, we can look closer at which VMs best fit your use case. Let's start by going over the types of VMs available in Azure Portal. You can also access the Microsoft documentation that includes all of these VMs.

Type	Sizes	Descriptions
General purpose	B, Dsv3, Dv3, Dasv4, Dav4, DSv2, Dv2, Av2, DC, DCv2, Dv4, Dsv4, Ddv4, Ddsv4, Dv5, Dsv5, Ddv5, Ddsv5, Dasv5, Dadsv5	Balanced CPU-to-memory ratio. Ideal for testing and development, small to medium databases, and low to medium traffic web servers.
Compute optimized	F, Fs, Fsv2, FX	High CPU-to-memory ratio. Good for medium traffic web servers, network appliances, batch processes, and application servers.
Memory optimized	Esv3, Ev3, Easv4, Eav4, Ebdsv5, Ebsv5, Ev4, Esv4, Edv4, Edsv4, Ev5, Esv5, Edv5, Edsv5, Easv5, Eadsv5, Mv2, M, DSv2, Dv2	High memory-to-CPU ratio. Great for relational database servers, medium to large caches, and in-memory analytics.
Storage optimized	Lsv2, Lsv3, Lasv3	High disk throughput and IO ideal for Big Data, SQL, NoSQL databases, data warehousing, and large transactional databases.
GPU	NC, NCv2, NCv3, NCasT4_v3, ND, NDv2, NV, NVv3, NVv4, NDasrA100_v4, NDm_A100_v4	Specialized VMs targeted for heavy graphic rendering and video editing, as well as model training and inferencing (ND) with deep learning. Available with single or multiple GPUs.
High performance compute	HB, HBv2, HBv3, HC, H	Our fastest and most powerful CPU VMs with optional high-throughput network interfaces (RDMA).

VM Categories

Looking at category types can be helpful in narrowing down the list and saving time.

- **General purpose** - As the name suggests, these VMs are versatile and intended for smaller workloads.
- **Compute optimized** - Meant for heavy computing which can be better for Direct Query workloads.
- **Memory optimized** - Proper VM for large caches meaning that it will work best with lots of Import mode use cases.
- **Storage optimized** - Best for leveraging large data warehousing jobs.
- **GPU** - Very specialized use cases that include heavy graphics and video editing.
- **High performance compute** - These are the most powerful (and expensive!) VMs available, designed for real-time use cases.

After determining which general category fits your use cases, you can decide what model of the VM fits the rest of your requirements. The best way to do this is to view the **VM selection screen** in your **Azure Portal**.

In your Portal, select **Virtual Machines** from **Azure Services**, click **Create**, then **Azure Virtual Machines**. About midway down the screen, you will see a section called **Instance Details** where you can choose a size for your VM. Click on the See all sizes button to view each of the SKUs with their associated prices.

Curious about Azure VM pricing? Check out the Azure VM pricing calculator for a cost estimate.

At the start of this section, it was mentioned that you will need to find out the size of your datasets to make the right decision regarding VM. Here is where that information comes into play.

Each VM SKU has different specifications in regard to CPU cores, RAM, disks, and temporary storage. These different specifications also come with a range of price tags. It is imperative to understand the different requirements to ensure that you select the appropriate VM for your Power BI deployment.

Microsoft has issued a recommendation for a VM with at least 8 CPU cores and 8 GB of RAM for a Power BI gateway. This guidance might be helpful as a starting point for excluding the VMs within your category that will not meet this criteria.

Once you've narrowed to the VMs that will meet the base requirements, you will need to identify a VM that meets your individual needs in regard to data size. If you are caching your data, you'll want a VM with an amount of RAM that meets - or preferably exceeds - your largest dataset sizes.

For example, if your post compression dataset sizes max out at 32 GB, you may want to choose a VM with 64 GB of RAM to make sure that the loads can be appropriately handled.

Finally, once you've found a few VMs that meet your requirements, you'll need to consider the price of getting the proper amount of VMs for your deployment. You will want to procure at least two of your specified VMs, since it is a deployment best practice to use at least two gateways to form a cluster.

Make sure that you're anticipating the cost of multiple VMs when you are taking the price into account.

Note: As you can see, sizing VMs can be a difficult exercise because of the complexity and number of variables involved in the process. Although not an exact science, we hope this white paper has helped answer some of your questions and removed some of the uncertainty around choosing the correct VM.

SETTING UP GATEWAYS ON AZURE VIRTUAL MACHINES

Once you've decided that you need gateways and you have selected the appropriate VMs to host them on, it's time to set up those gateways. The setup is very crucial to making sure that your data is secure and able to be refreshed properly on a premium capacity. Let's dive into how to do this.

HOW TO SET UP YOUR GATEWAYS ON A VM

To start, we'll need to go to the **Azure Portal** and select **Virtual Machines** from the services list. If you haven't yet selected your VMs, learn how to select the right VM for your deployment.
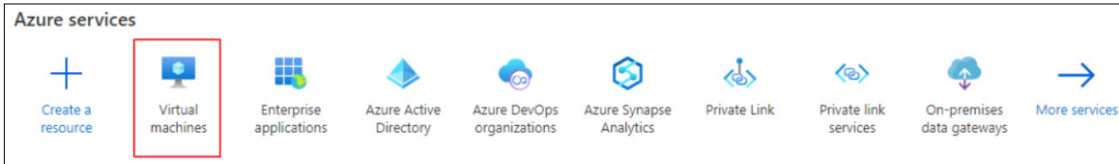
Next, you should be able to view all of the VMs that have been procured in Azure.
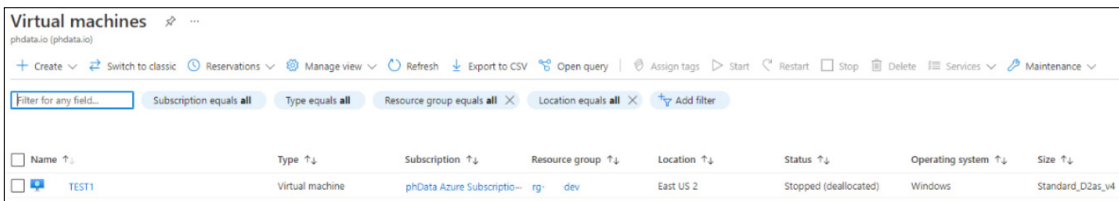
Then, you'll need to ensure that your VM is running. If it isn't, press the Start button. Once your VM is running, press the **Connect** drop-down and select **RDP** and then click **Download RDP File**.

Open the file and follow the prompts to log in to the VM. If you are doing this for testing and have access, you can log in as your normal user account. If you are intending to use this for enterprise deployment, you will need to use a service account to log in to the VM so that you can maintain continuity of the gateway despite changes in staffing.
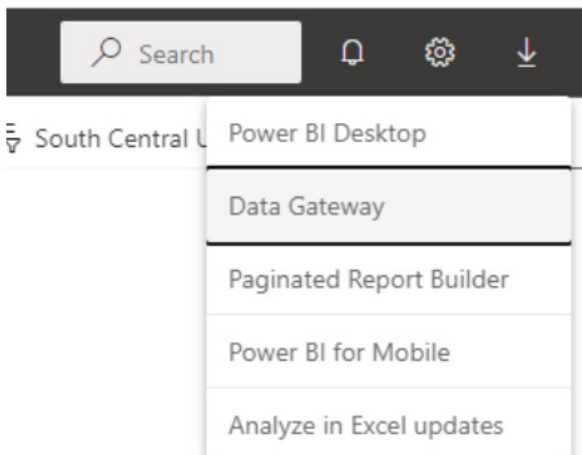
Once you have logged in, be sure to open an internet browser and go to PowerBI.com. Then click on the download icon and select **Data Gateway**.

Next, you should be prompted to go to Power BI Gateway where you can choose whether you're going to download using the Standard or Personal mode.

You are going to want to choose the Standard mode at this point for almost all enterprise deployments. If working only individually, you can use Personal mode, however most use cases we witness have some element of collaboration, which is why we suggest using Standard mode.

After you have downloaded the gateway with the appropriate mode to your VM, open the program. You should be able to see this popup. Ensure you are selecting the appropriate area to install the gateway. If you are using a VM, double check that you are using the correct drive that has all of the memory.
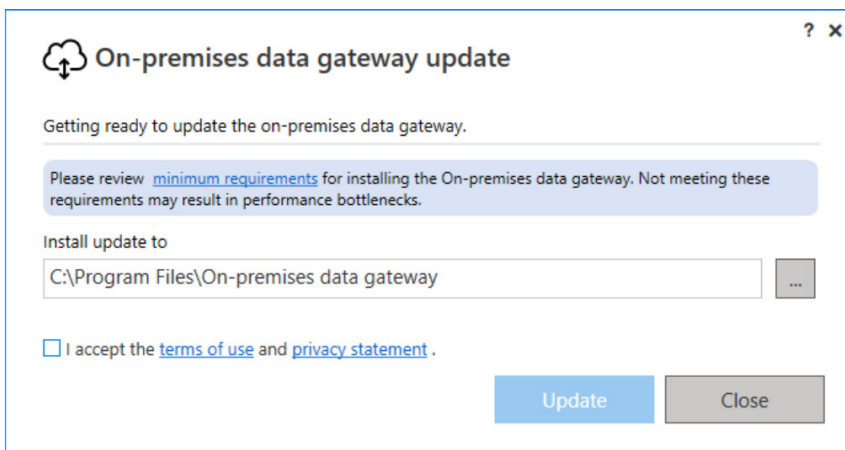
Once the gateway has been successfully installed to the appropriate machine and drive, you'll need to sign into the gateway using your service account. You should choose Register a New Gateway and then enter your new gateway name. Now your gateway will be ready for use.
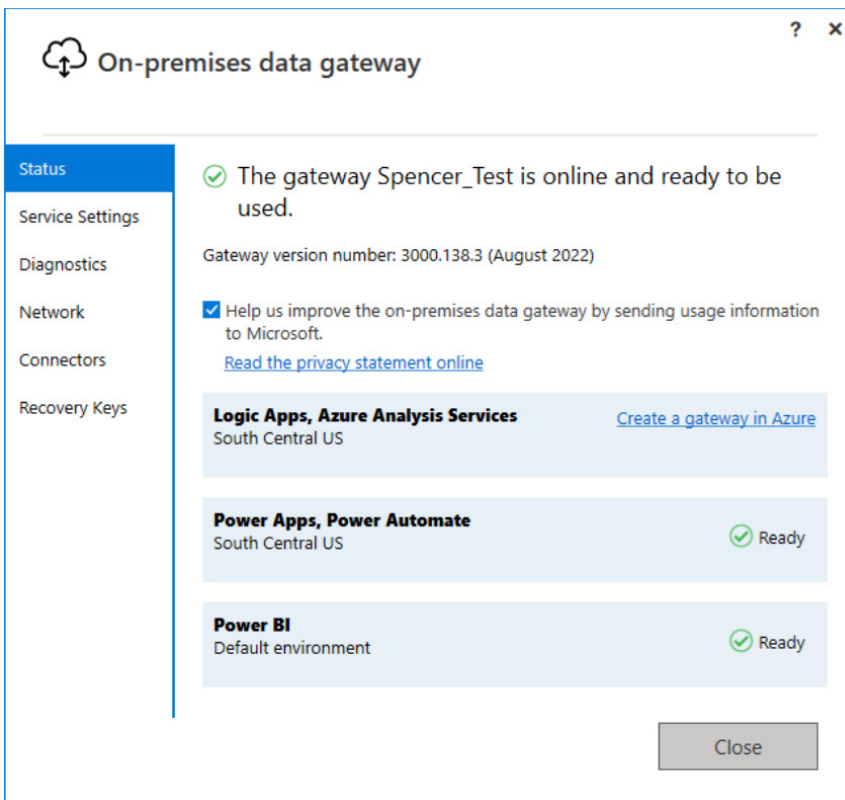
It's time to go to Power BI Service to complete the gateway setup. Navigate to the gear icon in the top right and choose **Manage Gateways**. Beforehand, make sure that you are able to view the gateway set up in the **On-premises data gateways** view.

Once you're certain you see the gateway in Power BI Service, click on **Manage Users** at the top to identify who else will be your gateway admins. You can either manage this via individual contributors or by Azure AD Group, which is the recommended format.

Next, a data source that utilizes the gateway will need to be added by going to the top left of the **Data Sources view** and selecting **+ New**. Then under **Gateway Cluster Name**, look for the name of the gateway that you just set up.

Make sure to identify the Server and Warehouse that you are utilizing if you are setting up a Snowflake data source. For the Authentication Method, select OAuth2 if you're setting up with SSO.

Make sure that the right credentials are submitted in the **Authentication** box by selecting **Edit Credentials**.

Note: These credentials should be those of your Service Account. The credentials will remain embedded in this gateway and be used to authenticate against your datasource. The users of the gateway will need to authenticate to the gateway, which we will highlight later in this whitepaper. This process will ensure that the gateway stays connected to the dataset and will not break even if an admin were to exit the company or change roles.
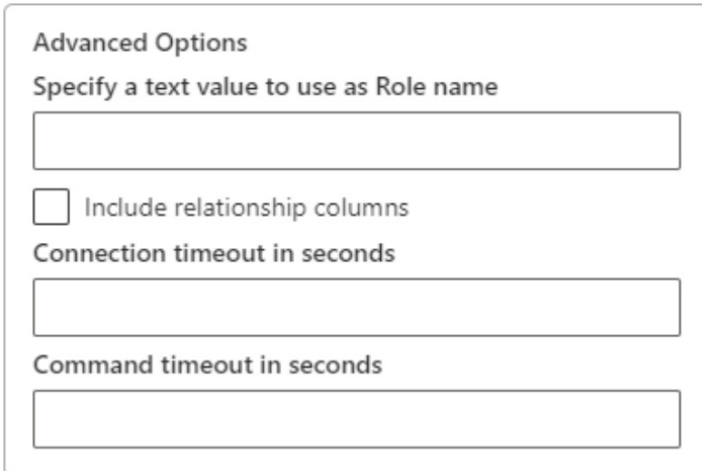
Moving down the menu, select the **Use SSO via Azure AD for DirectQueries** if you are wanting to enable that functionality. Finally, select **Organizational** for your enterprise set up in the Privacy Level box. Since there are multiple roles that you can use in Snowflake, if you do not specify a Role then it will default to the user's default role.

We're almost there!
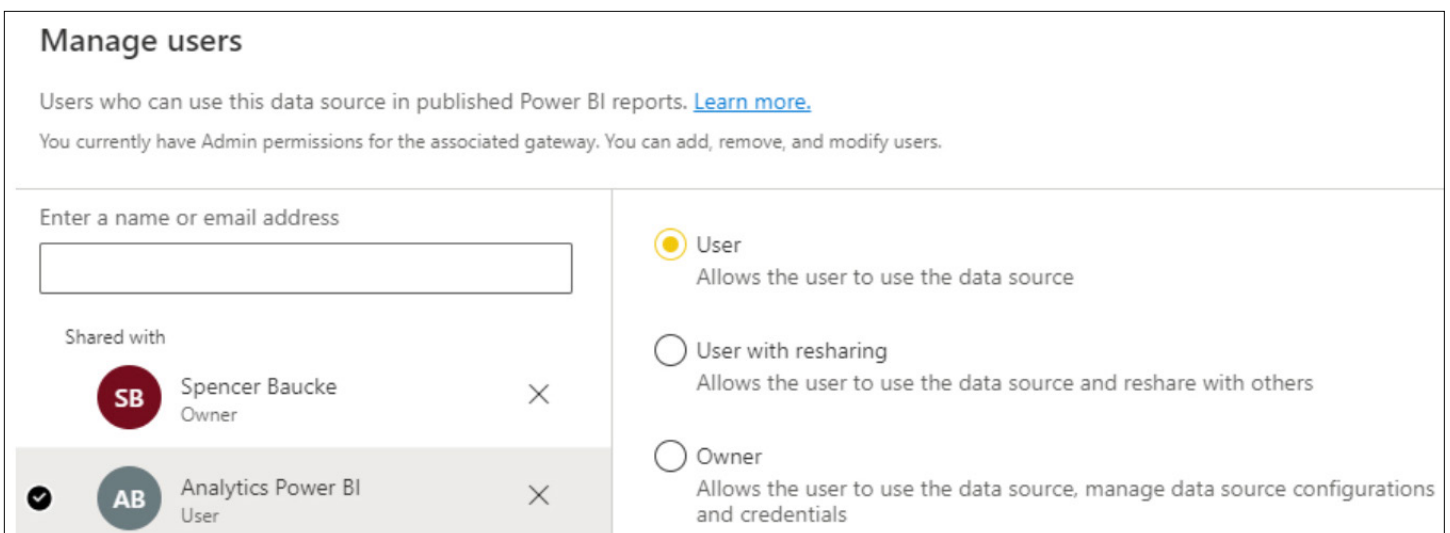
Now that you can view the gateway and the appropriate data sources have been set up, the last thing needed is to ensure that all desired users have access to the gateway. You can do this by choosing a data connection and then selecting **Manage Users** at the top of the screen.

Then enter in the user or user group that you want to have access to the data source, ensuring they have the appropriate intended level of access.
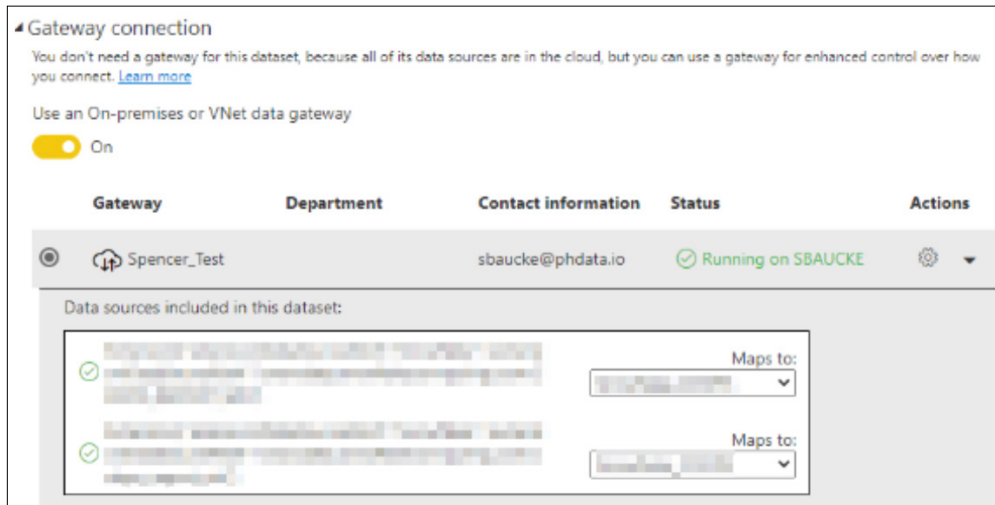
Congrats! Your data gateway has now been successfully set up! Now, it's time to ensure that your reports are utilizing the gateways appropriately.

SETTING UP GATEWAY WITH THE DATA SOURCE

Once your report and model have been published to Power BI Service, go to the data model and click the Schedule Refresh button so you can see the settings for the dataset. Next, go down to **Gateway Connection** and toggle the Use an On-Premise or VNet data gateway selector to **On**.

Double check that the Gateway maps to the data source you set up in the **Manage Gateways** portal.

Upon viewing the green checkmark in the **Status** section, your data source will be ready for a refresh from Power BI Service. To test this, go to the data set and click the **Refresh Now** button. Once you have completed the test and the data set is refreshing properly, you should have a properly functioning gateway and dataset!

SETTING UP AZURE AD SSO WITH SNOWFLAKE

The Snowflake Data Cloud and Power BI are currently two of the top technologies in the data landscape. Not surprisingly, many enterprises are using these two tools together. Because of their popularity, it is important for organizations to understand how these two tools play together.

WHY SNOWFLAKE AND POWER BI SPECIFICALLY?

Snowflake and Power BI are two of the most popular names in their respective fields. For several years now, Snowflake has been revolutionizing the cloud data industry. This robust platform showed huge growth and adoption resulting in the largest software IPO in history back in late 2020.

Likewise, Microsoft's Power BI has been making huge strides in the data analytics and visualization space. Much of this has been the result of the heavy investment that Microsoft has made in the tool which features monthly releases and an ongoing array of new features, e.g. Field Parameters, Datamarts, and the new Azure DevOps extension.

Many companies are leveraging these platforms in conjunction. Unfortunately, their collective popularity doesn't mean their integration is 100% seamless.

Because Power BI is a Microsoft product, it has a more seamless integration with tools like Azure Synapse Analytics. Despite this, we will help you make these two platforms work together with ease.

WHAT IS SSO, AND WHY DOES IT MATTER?

Single Sign On (SSO) is an authentication method that allows end users to authenticate into an application with one set of authentication credentials, instead of needing to authenticate into every database and dataset with individual credentials. A properly configured SSO enables one set of credentials to allow access to all of the necessary applications.

If you are considering a solution to deploy across your organization, SSO is the best way to ensure a seamless user experience. The most common scenario for using Snowflake and Power BI together is within an organization that already leverages the 365 suite, meaning they are most likely already leveraging Azure Active Directory groups for things like licensing and access.

Due to this, we are going to show how to set up SSO to Snowflake using Azure AD groups in Power BI.

STEPS TO ENABLE SSO ACROSS PLATFORMS

There are five main steps needed in order to achieve successful integration of Snowflake and Power BI.

1. Configure Azure for SSO on Snowflake
2. Configure Snowflake for SSO with Azure AD
3. Configure Power BI Service for SSO with Snowflake
4. Configure datasets with gateways on VMs
5. Load, Deploy, and Refresh datasets using gateways

How Snowflake and Power BI Communicate
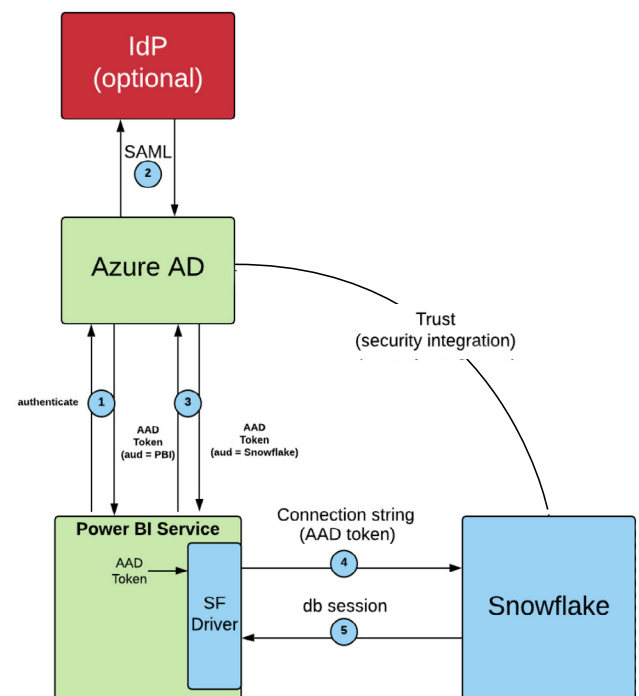
CONFIGURE AZURE FOR SSO ON SNOWFLAKE

Use these step-by-step instructions to set up SSO in Azure:

1. In Azure Portal, add Snowflake as a new application in your Enterprise Applications
2. Set up the Single Sign-On with SAML section
3. Assign Test User in the Users and Groups section of the Snowflake Application

CONFIGURE SNOWFLAKE FOR SSO WITH AZURE AD

Once you've set up the SAML component in Azure and assigned users or groups to the application, it's time to also set up Azure AD SSO on the Snowflake side. In Snowflake you will want to run the following script that references the token that is created in Azure. This will map usernames from Azure to Snowflake.

```
create security integration powerbi
 type = external_oauth
 enabled = true
 external_oauth_type = azure
 external_oauth_issuer = '<AZURE_AD_ISSUER>'
 external_oauth_jws_keys_url = 'https://login.windows.net/common/discovery/keys'
 external_oauth_audience_list = ('https://analysis.windows.net/powerbi/connector/Snowflake')
 external_oauth_token_user_mapping_claim = 'upn'
 external_oauth_snowflake_user_mapping_attribute = 'login_name'
 external_oauth_any_role_mode = 'ENABLE';
```

Snowflake users who are account administrators or have a role with the global create integration privilege can run this script only. Learn more about how to set up Azure AD SSO in Snowflake.

A common question asked about setting up Snowflake SSO is whether or not Azure Private Link is supported. The answer to this question is "YES, this is supported". Once SSO is set up normally, it should not be a big effort to get SSO set up with Private Link. More specific documentation is available from Snowflake.

CONFIGURE POWER BI SERVICE FOR SSO WITH SNOWFLAKE

Now that SSO capabilities are set up on both the Snowflake and Azure sides, we need to be sure that we have set things up in the Power BI portal appropriately to support this capability.

In your Power BI Tenant Settings, go down to the Snowflake SSO and Azure AD Single Sign-On for Gateway settings and enable these two settings to ensure:

1. The token information from Azure will be able to be sent to Snowflake

2. We can use Azure AD SSO while utilizing an on-premise gateway.

CONFIGURE GATEWAYS AND TEST

Now that Azure, Snowflake, and Power BI have been configured to accommodate SSO, you may need to install your on-premise gateway. A gateway is not necessary, but it is often recommended for security reasons for enterprise deployments, and for traversing a private network (Azure Private Link, as an example).

If you don't have a need for gateways, you can connect to Snowflake and test your Microsoft authentication. Be sure to be signed into your account in the Power BI Desktop. Also, use the Microsoft account selection when providing credentials.
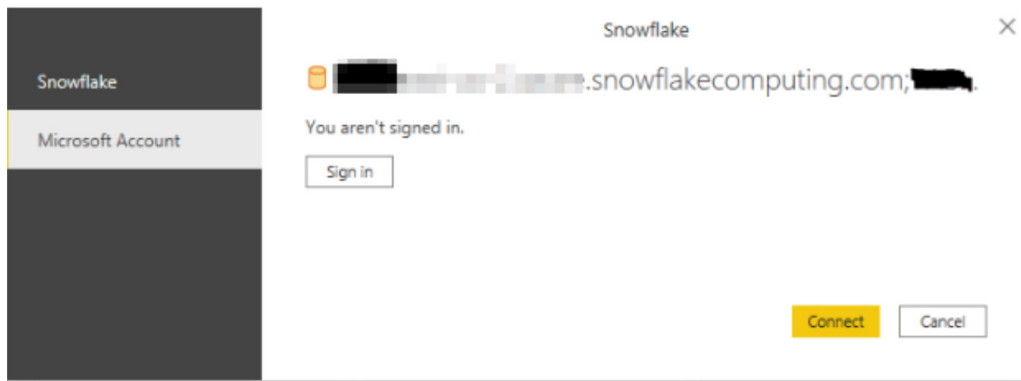
If the SSO has been improperly set up, you will receive an error message. The most common error message received for improper set up access is: "We couldn't authenticate with the credentials provided."

This error message means that somewhere along the chain of authentication, the token or username was not validated properly. In this scenario, check the setup on both the Azure and Snowflake sides then try again.

Learn more about connecting to Snowflake with Power BI.

TIP Be sure to involve your Azure and Snowflake admins in the set-up process because they may already have some of the infrastructure created. Because these admins will most likely be the ones setting up the scripts and configurations, make sure to coordinate with them to be sure that you are following the proper company procedures.

POWER BI SERVICE ADMINISTRATION

REFRESHING DATASETS IN POWER BI SERVICE

Refreshing a dataset in Power BI is relevant for imported datasets (as opposed to DirectQuery mode or a Live Connection). When using import mode, Power BI imports the data from its original data sources and stores it in the Power BI dataset. There are two methods of refreshing a dataset from Power BI Service. Datasets can either be refreshed on-demand or refreshed on a specified schedule.

Refreshing a dataset on demand is as easy as clicking the **Refresh now** button which appears next to the dataset in the workspace. A rotating semicircle of dots will appear in the Refreshed column indicating that the refresh is in process.

Users can also schedule datasets to be refreshed on a specified cadence that is determined in the dataset settings page. Datasets residing on a Premium capacity can be scheduled to refresh up to 48 times per day. To set the refresh schedule for a dataset, navigate to that dataset's settings, toggle the setting to **On** and make the appropriate selections for the following:

- **Refresh schedule:** if Daily you will select the hours the dataset refreshes each day. If a Weekly cadence is set up you will select the days of the week and hours of those days that the dataset will refresh.

- **Time zone:** determines what time zone the refresh schedule is associated with.

- **Refresh failure audience:** in case of a refresh failure, this determines who will receive a notification.

SETTING UP POWER BI DEPLOYMENT PIPELINES

The Power BI Deployment Pipelines tool is a feature within Power BI Service that enables developers and admins to manage the development lifecycle of content that their team is producing. The tool is designed to be an efficient and reusable pipeline that automates the movement of content through three stages:

1. **Development:** design, review, and revise content in a development workspace. When it's ready to be tested, deploy the content to the test stage.

2. **Test:** perform UAT and verify that the content is accurate in this preproduction workspace. After UAT has been completed and the content accepted, deploy it to the production stage.

3. **Production:** the workspace content has been tested and is ready to be consumed by your users in the workspace app.

Content in this context refers to reports, paginated reports, dashboards, datasets, dataflows, and datamarts.

Note that the different stages of a pipeline each have their own associated workspace and that it is not possible to assign multiple workspaces to one stage of the pipeline. The benefit of this is so that workspace admins can provision the proper access accordingly for each stage of the development lifecycle – for example, you may not necessarily want someone who is assigned as a "Viewer" to have access to the development or test workspaces.

To create a deployment pipeline, simply navigate to the **Deployment Pipelines** page using the sidebar in Power BI Service. You will be brought to a screen with a yellow **Create a pipeline** button followed by a popup asking for the name and description of the pipeline - provide that information.

The next window requires you to assign the appropriate workspaces to the three stages of the deployment pipeline.
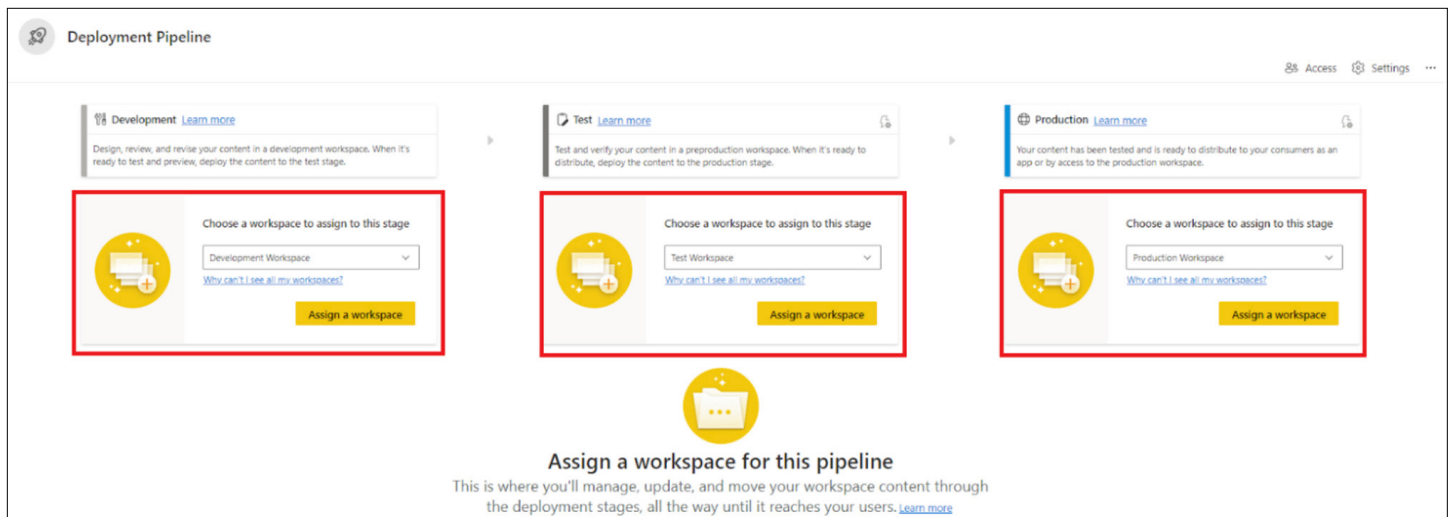
After completing the previous step, the deployment pipeline is now ready for use. You can view the content in the workspace associated with each stage of the pipeline by clicking the **Show more** drop down.

There are two ways to deploy content to the next stage in the pipeline, you can deploy everything in a workspace, or deploy only specific items in the workspace. You can also publish and update the app from this view by using the associated buttons in pipeline stages.

- To deploy the entire workspace, simply click the yellow **Deploy** button.
- To deploy specific content in a workspace, select the check box next to the content and then click **Deploy**.
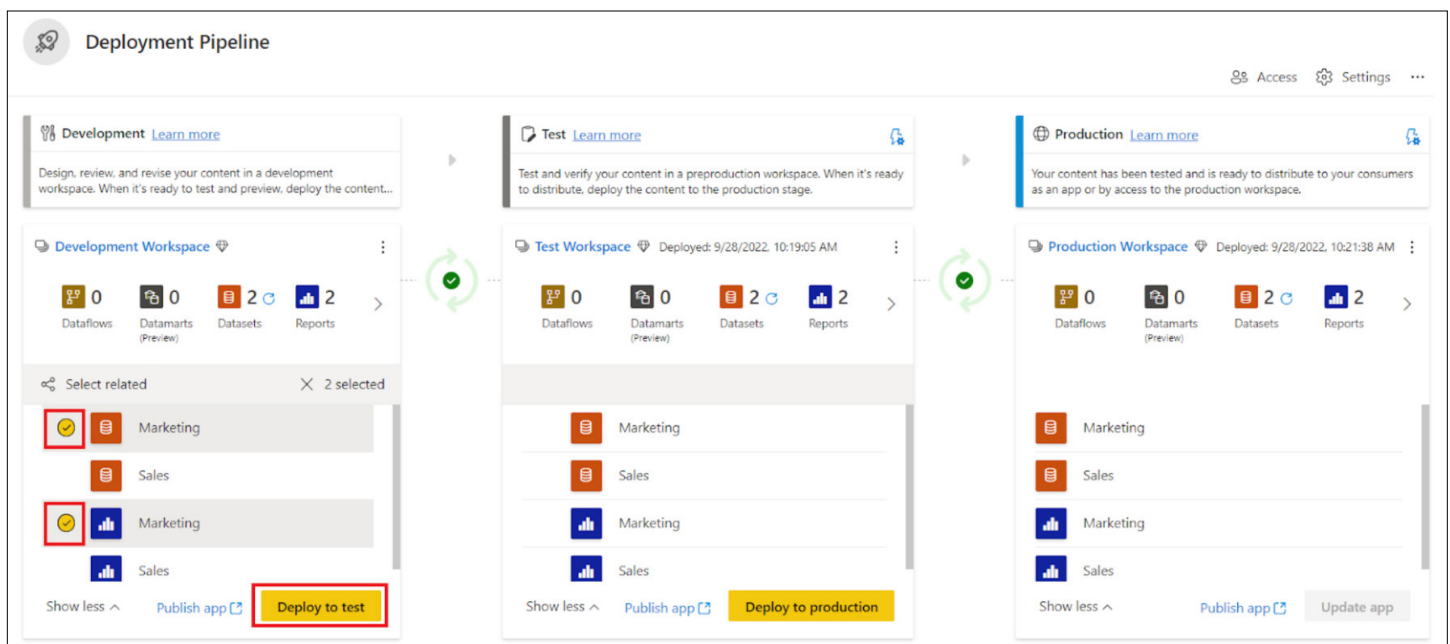
In between each stage of the pipeline you will notice either a green checkmark or an orange "X". These icons provide an indication that something may be asynchronous among the workspaces.

A green checkmark – all of the content in the two workspaces are a one-for-one match. Everything from the preceding workspace has been promoted to the next stage.

An orange "X" – a change has occurred and that the content needs to be promoted to the next stage to reflect the latest version. If this is the case, a user can click on the **Compare** button to view what is different from one workspace to another. In the example below, the Marketing report has been updated in the Test workspace but not the Production workspace.
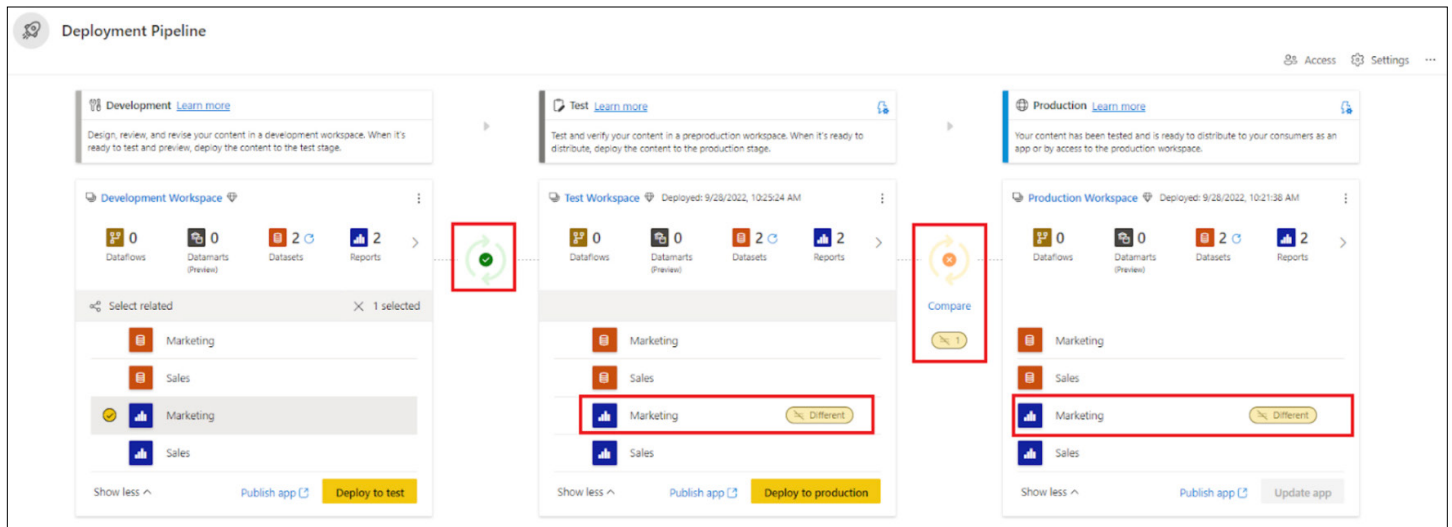
Power BI Deployment Pipelines also offers the ability to automate the task of replacing a dataset's data source. Teams will often use development or test data in the development lifecycle, this is done so that resources in the production data environment are reserved for production reports. To enable this functionality:

1. Click the blue lightning bolt next to the name of the Test or Production workspace.

2. In the side panel that appears, select the dataset for which the deployment rule is being created and then click the **+ Add rule** button.

3. Select the data source to be replaced and input the data source that it is being replaced with and click **Save**.

Now each time the dataset is deployed to that stage, the data source will be replaced with the new one.
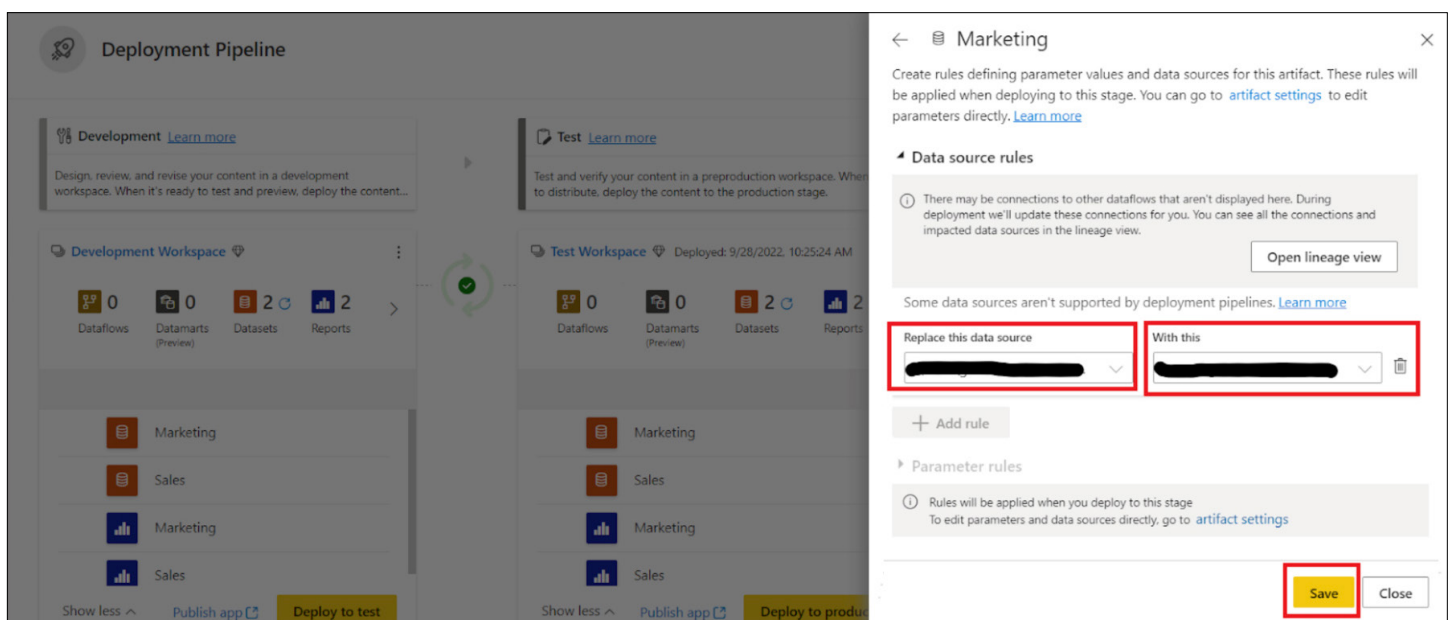
Access to a Power BI Deployment Pipeline and its associated workspaces are independent of each other – meaning that if someone has access to a deployment pipeline they don't necessarily have access to the workspaces and vice versa. Provision access to the pipeline by clicking the **Access** link on the top right of the Deployment Pipeline window.

USAGE METRICS REPORTING

Having a solid understanding of how reports in a workspace are used can be helpful in prioritizing development efforts. Fortunately, Power BI has an out of the box report that can be generated that provides usage metrics for a workspace. To create the report select the three dots next to the report title in the workspace and click **View usage metrics report**.
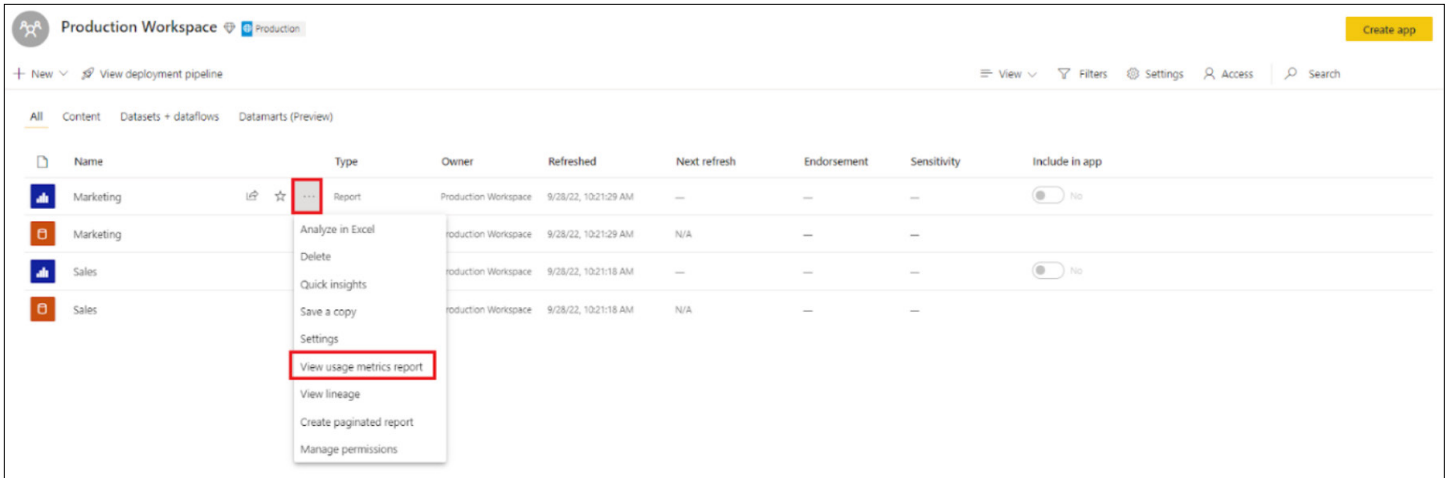
At the time of writing this whitepaper the workspace view is in Preview. The default report generated from the previous step will show metrics for just the report selected. To get the workspace view toggle the **New usage report on** switch on the top right.

The new usage report will have four tabs available - the data in these tabs is updated daily and shows the last 30 days of data.

- **Report usage:** shows basic details about report viewership including number of views, users of the report, and a trend analysis for the reports in the workspace that are filtered using the Filters panel.
- **Report performance:** details load times for reports.
- **Report list:** provides a list with all of the reports in a workspace along with associated metrics pertaining to usage.
- **FAQ:** used to gain a deeper understanding of how to use the Usage Metrics Report
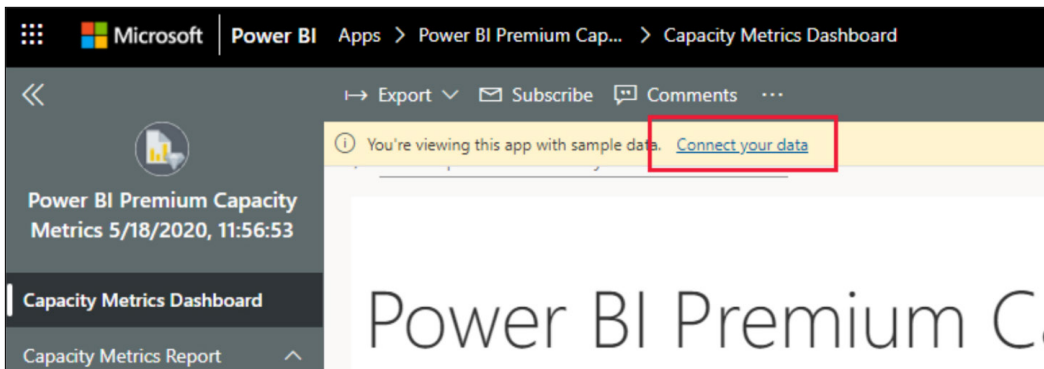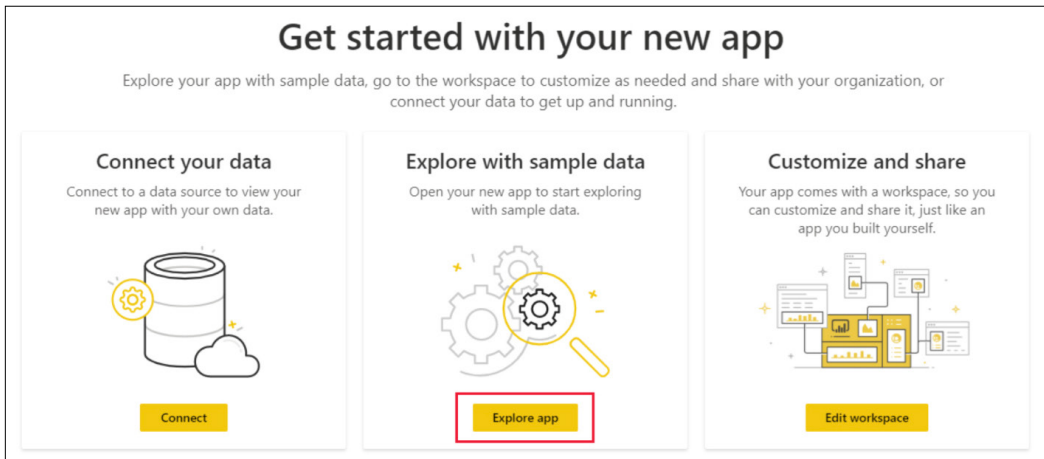
MONITORING PREMIUM CAPACITY REPORTING

Monitoring the overall Premium Capacity that you have set up gives administrators an understanding of how the capacity is performing and can assist in making informed decisions on how best to utilize the Premium Capacity resources. A Power BI Premium Capacity Metrics app is available for installation as a template app and provides in-depth details on how the capacity is performing.

To deploy the app follow these simple steps:

1. Install the app by clicking the following link: Power BI Premium Metrics template app and then select *GET IT NOW* followed by the yellow *Install* button.
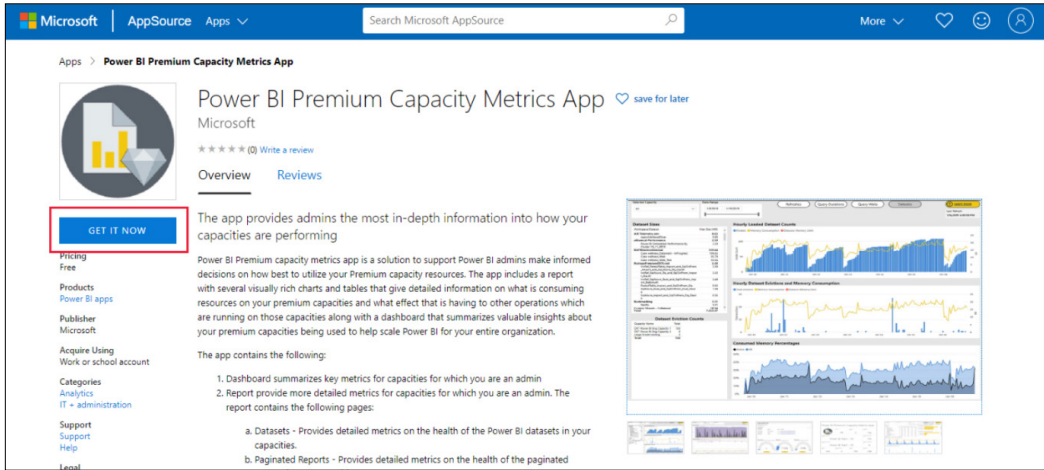
2. Once installed, the app will be available in the Apps page - the next step is to connect your data. Select the *Explore app* option and then *Connect your data* on the top of the app. This process may take a while.


Complete step-by-step directions are available from Microsoft. Connect to Power BI Premium Capacity Metrics

Monitoring Premium Capacity

After the data has been connected, the app is ready to be used to monitor your premium capacity. The app has seven different reports in their own tabs which can be navigated using the panel on the left.

The first report in the app is the **Capacity Health Center**.
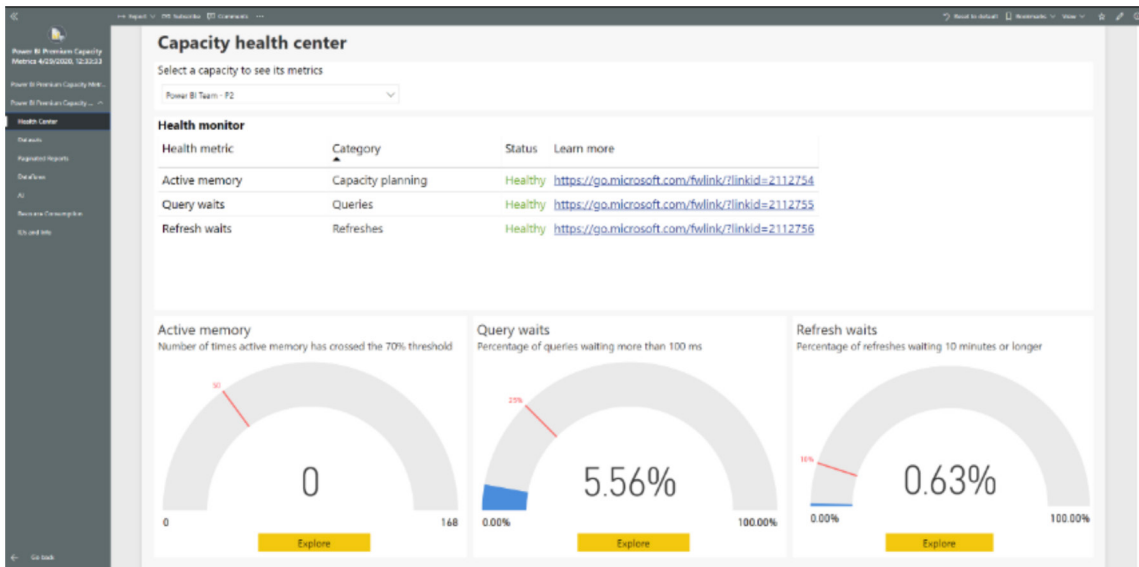
The Health Center provides an overview of the health of the capacity and has three metrics displayed using a gauge visualization that assist in doing so:

- **Active Memory:** this is a health indicator to evaluate the capacity resource consumption usage so that admins can adjust the capacity if necessary. The active memory is used to process active datasets and therefore will not be evicted when memory is needed. Use this to understand if your capacity can handle additional load, or if it's nearing or already over capacity. The visualization measures the number of times that the capacity's active memory has crossed the 70% threshold.

- **Query Waits:** this category indicates if users may experience report visual refreshes that are slower than normal or that could become unresponsive. The KPI measures whether 25% or more of the capacity's queries have been waiting 100 milliseconds or longer to execute. Query waits happen when the available CPU isn't enough to handle the pending queries.

- **Refresh Waits:** this metric is helpful in understanding if users may be experiencing stale data in their reports. Refresh waits is the time a given dataset refresh waited before starting to execute, from the time the dataset was triggered on demand, or from its scheduled refresh. The KPI measures whether 10% or more of pending refresh requests took 10 minutes or longer. Again, this generally occurs when there's an insufficient amount of CPU or available memory.

The remaining tabs are a deeper dive into the capacity. The Datasets, Paginated Reports, Dataflows, and AI tabs in the app show reports regarding usage metrics, refresh results, and memory consumption pertaining to the respective subjects. The Resource Consumption report provides details on CPU and memory consumption for the capacity. Finally, the IDs and Info report gives details about the capacity SKU and its admins.

For a detailed description of every single metric available in the Premium Capacity Metrics App, view the Microsoft Documentation here: [Monitor Premium capacities with the app](#).

IN CLOSING

Choosing to deploy Power BI Premium capacity is a big decision for you and your organization. At phData, our team of experts are ready to help you feel more comfortable with your choice and see the return on your investment in Power BI.

If you are planning a Power BI Premium capacity implementation or have already invested in Premium capacity but need help optimizing it, we would be more than happy to help you out!

In addition to our implementation services, our 24x5 Power BI support gives you more flexibility and peace of mind if needed.

EXPLORE PHDATA'S POWER BI SERVICES